US-PAT-NO: 6608930

DOCUMENT-IDENTIFIER: US 6608930 B1

TITLE: Method and system for analyzing video content using detected text in video frames

--------- KWIC ---------

Abstract Text - ABTX (1):

There is disclosed, for use in video text analysis system, a video processing device for searching **video streams** for one or more user-selected image text attributes. The video processing device comprises an image processor capable detecting and extracting image text from video frames, determining attributes of the extracted image text, comparing the extracted image text attributes and the user-selected image text attributes, and, if a match occurs, modifying, transferring, and/or labeling at least a portion of the **video stream** in accordance with user commands. The invention uses the user-selected image **text attributes to search** through an archive of video clips to 1) locate particular types of events, such as news programs or sports events; 2) locate programs featuring particular persons or groups; 3) locate programs by name; 4) save or remove all or some commercials, and to otherwise sort, edit, and save all of, or portions of, video clips according to image text that appears in the frames of the video clips.

Application Filing Date - AD (1):

19990809

Brief Summary Text - BSTX (2):

The present invention is directed, in general, to video processing systems and, more specifically, to a system for analyzing and characterizing a **video stream** based on the attributes of text detected in the content of the video.

Brief Summary Text - BSTX (6):

The detection of advertising commercials in a **video stream** is an also active research area. See R. Lienhart et al., "On the Detection and Recognition of Television Commercials," Proceedings of IEEE International Conference on Multimedia Computing and Systems, pp. 509-516 (1997); and T. McGee et al., "Parsing TV Programs for Identification and Removal of Non-Story Segments," SPIE Conference on Storage and Retrieval in Image and Video Databases, San Jose (Jan. 1999).

Brief Summary Text - BSTX (15):

To address the above-discussed deficiencies of the prior art, the present invention discloses a video processing device for searching or filtering **video**

streams for one or more user-selected image text attributes. Generally, "searching" **video streams** refers to searching in response to user-defined inputs, whereas "filtering" generally refers to an automated process that requires little or no user input. However, in the disclosure, "searching" and "filtering" may be used interchangeably. An image processor detects and extracts image text from frames in video clips, determines the relevant attributes of the extracted image text, and compares the extracted image text attributes and the user-selected image text attributes. If a match occurs, the video processing device may modify, transfer, label or otherwise identify at least a portion of the **video stream** in accordance with user commands. The video processing device uses the user-selected image **text attributes to search** through an archive of video clips to 1) locate particular types of events, such as news programs or sports events; 2) locate programs featuring particular persons or groups; 3) locate programs by name; 4) save or remove all or some commercials, and to otherwise sort, edit, and save all of, or portions of, video clips according to image text that appears in the frames of the video clips.

Brief Summary Text - BSTX (16):

It is a primary object of the present invention to provide, for use in a system capable of analyzing image text in video frames, a video processing device capable of searching and/or filtering **video streams** in response to receipt of at least one selected image text attribute. In an exemplary embodiment, the video processing device comprises an image processor capable of receiving a first **video stream** comprising a plurality of video frames, detecting and extracting image text from the plurality of video frames, determining at least one attribute of the extracted image text, comparing the at least one extracted image text attribute and the at least one selected image text attribute, and, in response to a match between the at least one extracted image text attribute and the at least one selected image text attribute, at least one of: 1) modifying at least a portion of the first **video stream** in accordance with a first user command; 2) transferring at least a portion of the first **video stream** in accordance with a second user command; and 3) labeling at least a portion of the first **video stream** in accordance with a third user command.

Detailed Description Text - DETX (46):

Image processor 120 subsequently stores the extracted frame text in image text work space 132 and the process continues with the next frame at process step 205. The sequence continues until text has been extracted from all frames of the selected video clip. Once video image text has been recognized and extracted, image processor 120 may further isolate and analyze the extracted text based upon system-selected or user-selected text attributes. Video sequences or clips may be searched or indexed based upon the text present in a frame or a group of frames. For example, the presence of the word "hotel" in a frame may indicate a high probability of a hotel scene in the present and following frames. The presence of street names and numbers indicates the possibility of city scenes. Identification of billboard **text can be helpful when searching** for frames with highway scenes.

Detailed Description Text - DETX (47):
It is also possible to use text for analyzing and labeling specific video frames. For example, image processor 120 may identify and extract live news reports that were broadcast as part of a particular news program or were given by a particular news reporter. Image processor 120 may do this by **searching for keywords in image text,** including "LIVE", the program name (e.g., "NEWS4 AT NITE"), or the reporter's name (e.g. "Carol Jones").

Claims Text - CLTX (1):
1. For use in a system capable of analyzing image text in video frames, a video processing device capable of one of filtering and searching **video streams** in response to receipt of at least one selected image text attribute, wherein said selected image text attribute is a non-character attribute, said video processing device comprising: an image processor capable of receiving a first **video stream** comprising a plurality of video frames, detecting and extracting image text from said plurality of video frames, determining at least one non-character attribute of said extracted image text, comparing said at least one extracted image text non-character attribute and said at least one selected image text attribute, and, in response to a match between said at least one extracted image text non-character attribute and said at least one selected image text attribute, at least one of: modifying at least a portion of said first **video stream**; transferring at least a portion of said first **video stream**; and labeling at least a portion of said first **video stream**.

Claims Text - CLTX (9):
9. An image text analysis system comprising: a video processing device capable of one of searching and filtering **video streams** in response to receipt of at least one selected image text attribute, wherein said selected image text attribute is a non-character attribute, said video processing device comprising: an image processor capable of receiving a first **video stream** comprising a plurality of video frames, detecting and extracting image text from said plurality of video frames, determining at least one non-character attribute of said extracted image text, comparing said at least one extracted image text non-character attribute and said at least one selected image text attribute, and, in response to a match between said at least one extracted image text non-character attribute.and said at least one selected image text attribute, at least one of: modifying at least a portion of said first **video stream**; transferring at least a portion of said first **video stream**; and labeling at least a portion of said first **video stream**; a display monitor for displaying said at least a portion of said first **video stream**; and a user input device for receiving user commands.

Claims Text - CLTX (17):
17. For use in a system capable of analyzing image text in video frames, a method of one of searching and filtering **video streams** in response to receipt of at least one selected image text attribute, wherein the selected image text attribute is a non-character attribute, the method comprising the steps of: receiving a first **video stream** comprising a plurality of video frames; detecting and extracting image text from the plurality of video frames; determining at least one non-character attribute of the extracted image text;

comparing the at least one extracted image text non-character attribute and the at least one selected image text attribute; and in response to a match between the at least one extracted image text non-character attribute and the at least one selected image text attribute, at least one of: modifying at least a portion of the first **video stream**; transferring at least a portion of the first **video stream**; and labeling at least a portion of the first **video stream**.

Claims Text - CLTX (21):

21. For use in a system capable of analyzing image text in video frames, computer executable process steps stored on a computer readable storage medium for performing one of searching and filtering **video streams** in response to receipt of at least one selected image text attribute, wherein the selected image text attribute is a non-character attribute, the computer executable process steps comprising the steps of: receiving a first **video stream** comprising a plurality of video frames; detecting and extracting image text from the plurality of video frames; determining at least one non-character attribute of the extracted image text; comparing the at least one extracted image text non-character attribute and the at least one selected image text attribute; and in response to a match between the at least one extracted image text non-character attribute and the at least one selected image text attribute, at least one of: modifying at least a portion of the first **video stream**; transferring at least a portion of the first **video stream**; and labeling at least a portion of the first **video stream**.